



Consumer and
Corporate Affairs Canada

Consommation
et Corporations Canada

(11) (A) No. **1 181 856**

(45) ISSUED 850129

(52) CLASS 354-49

(51) INT. CL. G10L 1/04, G06G 7/60³

(19) (CA) CANADIAN PATENT (12)

(54) Personal Access Control System Using Speech and Face
Recognition

(72) Feix, Wolfgang H.,
U.S.A.
Ruell, Hartwig E.,
Germany (Federal Republic of)

(73) Granted to Siemens Aktiengesellschaft
Germany (Federal Republic of)

(21) APPLICATION No. 415,979

(22) FILED 821119

(30) PRIORITY DATE U.S.A. (323,657) 811120

NO. OF CLAIMS 13

Canada

DISTRIBUTED BY THE PATENT OFFICE, OTTAWA.
CCA-274 (11-82)

ABSTRACT OF THE DISCLOSURE

A method and an apparatus are disclosed for identifying an individual through a combination of both speech and face recognition. The voice signature of an interrogated person uttering a key word into a microphone is compared in a pattern matcher with the previously stored voice signature of a known person uttering the same key word to obtain a first similarity score. At the same time, when a key event in the utterance of the key word by the interrogated person occurs, a momentary image of that person's mouth region onto which a grid pattern has been projected is optically recorded and compared with the previously stored corresponding momentary image of the same known person to obtain a second similarity score. The two similarity scores are analyzed to verify that the identity of the interrogated person is that of the known person.

1

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to personal access control systems in general and, in particular, to a method and an apparatus for identifying an individual through a combination of speech and face recognition.

2. Description of the Prior Art

10 Speech recognition methods and apparatus have been used extensively in personal access control systems to limit access to secure facilities and to prevent the unauthorized use of information input and output devices of computers and various other machines. These systems 15 analyze voice input signals to determine the identity or non-identity of an individual who is seeking access to the facility or use of the device.

In a typical system of this type, the individual 20 seeking access or use is requested to utter a particular key word from among a sequence of predefined key words. The utterance of the key word is detected and analyzed by the speech recognition apparatus. The detected voice signature of the uttered key word is 25 compared to a predetermined stored voice signature corresponding to the utterance of the same key word by a previously cleared known individual. Access is permitted when the compared voice signatures of the uttered key word and the stored key word are sufficiently similar to indicate identity of the individual seeking access with the known individual. An example 30 of such a speech recognition system is described in



1 U.S. Patent 4,239,936, entitled "Speech Recognition System", which issued December 16, 1980.

Personal identification using such speech recognition systems can be sufficiently accurate and reliable only if an indefinite computing time is available in which to analyze the uttered key word. But to avoid unacceptable waiting time, in practice the recognition process must be completed within a period of time of about three seconds or so from the initial request for access. For this shortened operation time, personal access control using speech recognition alone is subjected to identification error (the wrong individual is cleared or the right individual is not cleared) and falsification (voice impression, tape recordings, etc.). Further, because of the difficulty of detecting the beginning and duration of speech signals corresponding to utterance of the key word, current speech recognition systems must use highly sophisticated technology, including costly speech signal duration detecting units. Moreover, it has been found that an increase in technical effort to achieve higher speech recognition system accuracy does not produce a proportional increase in the detection accuracy.

Personal access control systems have also been implemented using visual recognition for identification of individuals. Visual recognition systems use characteristic portions of the human body for identification purposes. Typical of this type of access control are fingerprint recognition systems and facial feature recognition systems. One such system is described in

1181856

-4-

- 1 U.S. Patent 4,109,237, entitled "Apparatus and Method for Identifying Individuals through the Retinal Vasculature Patterns", issued August 22, 1978. This latter system uses a method of scanning the individual's eye
- 5 with a light source arranged in a selected pattern and detecting that portion of the light source pattern which is reflected from the person's retina, thereby locating each intercept of the light source pattern with a blood vessel. The intercept pattern thus
- 10 obtained is then compared with stored intercept patterns previously obtained from individuals who are cleared for access. Personal access control systems using visual recognition alone demand an even higher level of technical effort and sophistication than
- 15 acoustical recognition systems.

-5-

1

SUMMARY OF THE INVENTION

It is an object of the present invention to provide
a method and an apparatus for identifying an individual
5 through a combination of both speech and face recogni-
tion which alleviates the disadvantages of and pro-
vides greater identification accuracy than personal
access control systems using either speech recognition
or voice recognition alone.

10

The method of the present invention provides for identi-
fying an individual through a combination of speech
and face recognition as follows: A characteristic
sequence of features of the voice is defined in
15 response to the utterance of a predetermined key word
by the individual to be identified. A momentary image
of a voice-utterance varying portion of the individual's
face is formed upon the occurrence of a key event in
the utterance of the key word. The defined sequence
20 of voice features and the momentary image of the
facial portion are then both used to determine the
identity or non-identity of the individual.

10

In a preferred embodiment of the method of the invention,
25 described in detail below, a first similarity ^{Score} rate is
computed by comparing the characteristic sequence of
voice features defined in response to utterance of the
predetermined key word by the individual by means of
a pattern matcher with a stored reference sequence of
30 features previously obtained from utterance of the
key word by a known person. When a key event in the
utterance of the key word by the individual occurs, the

1 momentary image corresponding to the moment of occur-
rence of the key event is stored. A second similarity
~~rate~~ is computed by comparing the stored momentary
image thus obtained with a stored reference momentary
5 image.

The ~~score~~
A second similarity ~~rate~~ is computed by comparing the
momentary image of the voice-utterance varying portion
of the individual's face corresponding to the moment
10 of occurrence of a key event in the utterance of the
key word with a stored, previously obtained reference
momentary image corresponding to the key event in the
utterance of the key word by the known person. Iden-
tity of the interrogated individual with the known
15 individual is determined when the first and second
similarity ~~scores~~ are above preselected coincidence
thresholds.

The apparatus according to the invention includes
20 means for defining a characteristic sequence of features
of the voice in response to the utterance of a pre-
determined key word by the individual to be identified
and means for forming a momentary image of a voice-
utterance varying portion of the individual's face
25 upon the occurrence of a key event in the utterance
of the key word. Connected to both the voice feature
sequence defining means and the momentary image
forming means are identification means for using both
the defined sequence and the momentary image to deter-
30 mine the identity or non-identity of the individual.

In a preferred embodiment of the apparatus, detailed
below, the voice feature sequence defining means

comprises a microphone, a preamplifier and an extractor. The momentary image forming means
1 comprises a camera, a detector, a memory and a key event detecting unit. The identification means connected to both the defining means and the momentary image forming means includes first and second
5 pattern matchers; first, second and third buffers; a microprocessor control unit and communicating means.

The method and apparatus of the present invention permits the realization of an efficient hybrid personal
10 access control system using a combination of both speech and face recognition. The invention offers improved performance over existing devices, with greater identification accuracy and security protection. Because both speech and face recognition techniques
15 are provided, identification accuracy at specific speech comparison thresholds and facial feature comparison thresholds is greater than for the same thresholds using only one of those techniques.

20 There have thus been outlined rather broadly the more important objects, features and advantages of the invention in order that the detailed description thereof follows may be better understood, and in order that the present contribution to the art may be better
25 appreciated. There are, of course, additional features of the invention that will be described more fully hereinafter. Those skilled in the art will appreciate that the conception on which this disclosure is based may readily be utilized as the basis for the designing
30 of other arrangements for carrying out the purposes of this invention. It is important, therefore, that

1181856

-8-

- 1 this disclosure be regarded as including such equivalent arrangements as do not depart from the spirit and scope of the invention.

1 BRIEF DESCRIPTION OF THE DRAWINGS

5 Embodiments of the method and apparatus of the present invention have been chosen for purposes of illustration and description. The embodiment of the apparatus which utilizes the method is shown in the accompanying drawings forming a part of the specification, wherein:

10 Fig. 1 is a block diagram of the apparatus of a personal access control system in accordance with the present invention;

15 Fig. 2 is a more detailed diagram of part of Fig. 1;

20 Figs. 3-5 are schematic representations of an individual uttering a key word which are helpful in understanding the image forming operation of the apparatus of Fig. 1; and

25 Fig. 6 is a graphical representation of the speech signal energy vs. time for the utterance of the key word by the individual in Figs. 3-5.

Throughout the drawings, like elements are referred
25 to by like numerals

1 DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Fig. 1 illustrates apparatus forming the basis of a personal access control system which identifies an individual based upon the vocal utterance by the individual of a specified key word. The speech or acoustical signal produced by the individual in the utterance of the word is detected and used to define a sequence of voice features. Simultaneously, certain facial features of the individual which vary when the key word is uttered are optically scanned and a momentary image is recorded of the physical position of the facial features at a prespecified time (a "key event") in the utterance of the key word. The sequence of voice features thus defined and the momentary image thus recorded are then both respectively compared to stored voice features and facial features previously developed from earlier vocal utterance of the same key word by a known individual. If there is sufficient coincidence of the "live" speech and facial features with the stored speech and facial features, the interrogated individual is cleared for access (i.e. the "identity" of the individual is determined). If there is not enough coincidence of both speech and facial features, the interrogated individual is not cleared for access (i.e. the "non-identity" of the individual is determined).

Referring to Fig. 1, the identification process is initiated when an individual requests access to a security zone or the like by dialing a certain personal identification number or by inserting a personal identification card into an input device, such as a

1 conventional key board 1. A microprocessor control
unit 2 such as an Intel SAB 8080 microprocessor
electrically connected for data communication with the
key board 1, receives the personal identification input
5 information from the key board 1. This input infor-
mation specifies the person whose identity is to be
verified. Responsive to receipt of this input, the
microprocessor control unit 2 communicates a pre-
determined key word to the individual to be interro-
10 gated by means of a display 3, such as a known LED-
display. The key word is determined by random selec-
tion from among a plurality of previously specified
key words which are stored in a memory 4 within the
microprocessor unit 2. At the same time, the control
15 unit 2 activates a microphone 5 which is coupled to
a preamplifier 6 and also activates a grid projector
7 which is associated with an electronic camera 8.
The grid projector 7 operates to project a grid pattern
onto a voice-utterance varying portion of the indi-
vidual's face. Such pattern may, for example, take
20 the form of the line pattern shown in Figs. 3-5, which
is projected onto the mouth region of the individual.
The grid projector 7 used to project the grid pattern
for identification purposes is in accordance with known
25 techniques, such as described in M. Fallah, "Biomedical
Imaging Processing for Dental Facial Abnormalities",
pages 462-464 (Department of Orthodontics, School of
Dental Medicine, University of Pittsburgh, Pittsburgh,
Pennsylvania).

Once the grid pattern has been projected onto the
individual's face, the electronic camera 8 focuses
on the mouth region of the individual and is acti-
vated to evaluate the distortions of the grid on the

1 mouth region. The camera 8 can be any suitable optical scanning device, such as a raster scanning camera, sensitive to visible and/or infrared regions of the electromagnetic spectrum.

5

When the individual utters the key word requested by the display 3, the individual's mouth region is being scanned by the electronic camera 8 operating at a standard TV camera scan frequency. Analog signals corresponding to a sequence of momentary images of the mouth region of the individual's face are thereby delivered to a detector 9, such as described in U.S. Patent 4,109,237. The detector 9 converts the analog signals of the camera 8 into digital signals, thereby creating a sequence of momentary images in the form of digital signals at the output of the detector 9.

As the individual speaks, the microphone 5 receives the acoustical voice signals and converts them by means of an associated preamplifier 6 into an electro-acoustical signal. The electro-acoustical signal is transmitted to a feature extractor 10. The feature extractor 10 performs a spectrum analysis of the input electro-acoustical signal and defines a characteristic sequence of features of the voice of the individual uttering the key word. This sequence of features is assembled into a voice signature of the interrogated individual. The voice signature can be a compilation of characteristic frequencies of the voice, or any other desired voice signature and is obtained by known techniques, such as described in U.S. Patent 4,239,936.

1181856

-13-

1 Connected to the feature extractor 10 is a pattern
matcher 11. The pattern matcher calculates the
measure of similarity between the "live" input voice
signature supplied by the feature extractor 10 and
5 a reference voice signature stored in a buffer 12.
The reference voice signature is entered into the
buffer 12 from the memory 4 in response to the iden-
tification process initiation and is the previously
10 stored voice signature for the uttered key word of
the person identified by the personal information
input number or card.

The electro-acoustical signals are simultaneously
delivered from the preamplifier 6 to a key event
15 detecting unit 13. The key event detecting unit 13
is connected to control a memory 14 coupled to the
detector 9, so that the memory stores the digital
signals of the momentary image in the sequence of
the momentary images delivered from the electronic
20 camera 8 which corresponds to the moment of occur-
rence of a key event described in the uttered key
word, as further described below.

The key event detecting unit 13 comprises an integra-
25 tor 15 connected to receive the electro-acoustical
signal from the preamplifier 6 in response to the
vocal utterance of the key word by the individual.
The integrator 15 operates to form a time dependent
signal corresponding to the energy of the electro-
30 acoustical signal. A representative time depen-
dent signal formed in response to utterance of the
key word is shown in Fig. 6. The integrator 15 may

take the form of a low pass filter to develop the time dependent signal in an analog way. Alternatively, as shown in Fig. 6, the time dependent signal may be developed in a digital way by sequentially deriving the square of the magnitude of the amplitudes of the electro-acoustical signal for successive intervals A₁, A₂..., A_N of about 10-20 milliseconds each, over a certain time period (called a "time window"). The time periods A₁, A₂..., A_N are overlapping, as shown in Fig. 6. The multiplications for the designated "time events" t₀ to t_n in Fig. 6 are used to define the shape of the signal energy. For each time event t₀ to t_n a different momentary image of the mouth region is detected (see Figs. 3-5). An 10 integrator of this type is within the skill of the art as described in U.S. Patent 4,109,237.

Coupled to the output of the integrator 15 is a control unit 16 which detects the beginning of a key word (t₀ in Fig. 6) by analyzing the output signal of the integrator 15. The control unit 16 corresponds to the "duration detecting unit" described in U.S. Patent 4,239,936. The beginning of a key word is detected by the control unit 16 by determining whether the amplitude of the signal is greater than the starting threshold (Fig. 6). Having detected the beginning of the key word, the control unit 16 activates a comparator 17 which is coupled to a slope detector 18 as well as to the control unit 16. The comparator 17 compares characteristic slope features of the energy signal (represented, for example, by the time events within a detecting time window) with previously stored slope features stored in a buffer 21 which define the key event and thereby detect the appearance of a key word. The characteristic slope features used to define the key event may be selected in many ways and the choice is largely a matter of individual preference. One way to define the key event is, for example, the moment of occurrence of a starting threshold of a certain magnitude followed by certain magnitudes of the signal energy at two specified successive 20

time events t_2 and t_4 within a preselected detecting time window. The key event is specified in terms of relative magnitudes of the threshold and amplitudes at t_2 and t_4 rather than in terms of absolute magnitudes which are subject to conditional variations. The circuitry needed for defining the key event in this manner is constructed using known techniques (such as using threshold detectors, counters, comparators and logic elements) and may be performed in either an analog or digital way.

When the occurrence of the key event has been detected by the comparator 17, a storing signal is delivered to the memory 14 causing the memory 14 to store the 10 momentary image of the mouth region corresponding to the key event. For example, the memory 14 may be directed to store the momentary image of the distorted grid pattern shown in Fig. 4 corresponding to the time event t_4 , in response to the detection of the threshold, amplitude at t_2 and amplitude at t_4 , all within the specified detecting time window. Connected to the memory 14 (controlled by the key event detecting unit 13) is a second pattern matcher 19 for computing a second similarity rate corresponding to the amount of similarity between the momentary image stored in the memory 14 and a reference momentary image stored in a buffer 20 coupled to the second pattern matcher 19 and to the microprocessor control unit 2. The reference momentary image is delivered to the buffer 20 from the memory 4 in response to initiation of the identification process and corresponds to the previously stored momentary image at the key event of the grid pattern projected onto and distorted by the mouth of the person specified by the input information in the utterance of the key word.

The buffers 12, 20 and 21 connected respectively to the first pattern matcher 11, the second pattern matcher 19 and the comparator 17 are all coupled for data communication to the microprocessor unit 2 by means of a data-bus line 22. Refer-

ence voice signatures, momentary images and energy signal characteristics (e.g. threshold and signal magnitude values) to define the key event corresponding to the utterance of each possible key word by a plurality of cleared, known persons are stored within the main memory 4 which is addressed by the microprocessor unit 2 for the chosen key word and person named by the input information. The buffers 12, 20 and 21 are loaded with comparison data according to the key word displayed to the individual on the display 4.

The first pattern matcher 11 and the second pattern matcher 19 are coupled to the microprocessor control unit 2 which includes a decision unit 23. The micro-
10 processor unit 2 compares the first similarity score computed by means of the first pattern matcher 11 and the second similarity score computed by means of the second pattern matcher 19 with acceptable predetermined

1 similarity ~~scores~~ stored in the memory 4 of the micro-
processor unit 2. If both the first and second simi-
larity ~~scores~~ exceed the preselected comparison rate
thresholds, the identity of the interrogated indi-
5 vidual with the person specified by the input number
or card is verified. If either ~~rate~~ score is below its
respective specified threshold, non-identity is
determined and access is denied. The result of the
evaluation process is shown on the display 3.

10

The design of the second pattern matcher 19 is shown
in Fig. 2. The second pattern matcher 19 comprises
an AND-gate 24 connected to an adder 25. The momen-
tary images stored in digital form in the memory 14
15 and the buffer 20 are retrieved by sequential address-
ing. The adder 25 counts whenever a coincidence occurs
between the reference signal 26 from the buffer 20 and
the momentary image signal 27 delivered from the memory
14. An additional AND-gate 28 connected to the output
20 of the adder 25 serves as a switch to deliver the
results of the matching process to the microprocessor
unit on request in response to a score signal 29
delivered from the microprocessor unit 2. The micro-
processor unit 2 also delivers an enabling signal to
25 the adder 25.

Having thus described the invention with particular
reference to the preferred forms of the method and
apparatus for a hybrid personal access control system
30 using both speech and face recognition techniques, it
will be obvious to those skilled in the art to which
the invention pertains, after understanding the invention,

-18-

1 that various changes and modifications may be made
therein without departing from the spirit and scope
of the invention as defined by the claims appended
hereto. For example, the choice of key words, the
5 characteristic sequence of features of the voice
selected for analysis, and the method of selection of
a key event to control storage of the "live" momentary
image are all matters of choice and can be varied to
suit individual preferences. Further, the use and
10 type of a grid pattern for projection onto an indi-
vidual's face is a matter of individual selection
and other optical scanning techniques can be used.
The choice of the grid pattern and mouth features as
described is made only as a convenient way to obtain
15 optical image comparison data of a voice-utterance
varying portion of the individual's face which can be
coordinated with information obtained from the
individual's speech in utterance of a preselected
word or preselected words. Optical scanning of the
20 eyes, nostrils, throat or cheeks also present possible
candidates for speech related examination as do the
lungs and other parts of the anatomy not normally
considered as part of the face. The term "voice-
uttering varying portion of the individual's face"
25 as used herein and in the claims is intended to be
defined broadly to encompass such other possibilities.

Additionally, while the personal access system described
in detail above is of an identification verification
30 type, those skilled in the art will appreciate that
the invention encompasses other systems, such as
systems which exclude certain individuals but permit
access to all others.

-19-

- 1 A personal access control system developed in accordance with the principles of the invention as defined above offers greater identification accuracy and reliability for the same complexity and sophistication of the
- 5 utilized apparatus than a system utilizing speech recognition or individual physical feature recognition, since with a system in accordance with the present invention, the simultaneous occurrence of two related identification parameters is being verified.

WHAT IS CLAIMED IS:

1. A method for identifying an individual through a combination of speech and face recognition which comprises:

5 a) defining a characteristic sequence of features of the voice in response to the utterance of a predetermined key word by the individual to be identified;

10 b) forming a momentary image of a voice-utterance varying portion of the individual's face upon the occurrence of a key event in the utterance of the key word; and

15 c) using both the defined sequence of features and the momentary image in order to determine the identity or non-identity of the individual.

2. A method for identifying an individual through a combination of speech and face recognition which comprises

5 a) defining a characteristic sequence of features of the voice in response to the utterance of a predetermined key word by the individual to be identified;

10 b) forming a sequence of momentary images of a voice-utterance varying portion of the individual's face upon the occurrence of a sequence of key events in the utterance of the key word; and

15 c) using both the defining sequence of features and the sequence of momentary images to determine the identity or non-identity of the individual.

3. A method according to claims 1 or 2, which further comprises communicating the predetermined key word to the individual in response to a request.

4. A method according to claims 1 or 2, which further comprises projecting a grid pattern onto the voice-utterance varying portion of the individual's face.

5. A method according to claim 1, wherein the momentary image forming step comprises:

5 a) scanning the voice-utterance varying portion of the individual's face with an imaging device and thereby creating a sequence of momentary images; and

10 b) detecting the occurrence of the key event and storing that momentary image in the sequence of momentary images which corresponds to the moment of occurrence of the key event;

15 and wherein the identity determining step comprises:

15 a) computing a first similarity ^{Score} rate between the defined sequence of the features and a reference sequence of features;

b) computing a second similarity ~~rate~~^{Score} between the stored momentary image and a known reference momentary image; and

20 c) determining the identity or non-identity of the individual by evaluating the computed ~~scores~~^{Score} first and second similarity rates.

6. Apparatus for identifying an individual through a combination of speech and face recognition which comprises:

5 a) means for defining a characteristic sequence of features of the voice in response to the utterance of a pre-determined key word by the individual to be identified;

10 b) means for forming a momentary image of a voice-utterance varying portion of the individual's face upon the occurrence of a key event in the utterance of the key word; and

15 c) identification means connected to both the defining means and the momentary image-forming means for using both the defined sequence of features and the momentary image in order to determine the identity or non-identity of the individual.

20 7. Apparatus according to claim 6, which further comprises:

means associated with the defining means
for communicating the predetermined key word
5 to the individual in response to a request.

8. Apparatus according to claims 6 or 7 wherein
the momentary image forming means comprises means for
projecting a grid pattern onto the voice-utterance
varying portion of the individual's face.

9. Apparatus according to claim 6,
a) wherein the momentary image-forming means
comprises:

5 means for scanning the voice-utterance varying
portion of the individual's face to create a
sequence of momentary images, and

10 means for detecting the occurrence of the
key event and storing the momentary image
in the sequence of momentary images which
corresponds to the moment of occurrence of
the key event; and

15 b) wherein the identification means comprises:
a first pattern matcher connected to compute a
first similarity ^{score} between the sequence of
features and a reference sequence of features, and

20 a second pattern matcher connected to compute a
second similarity ^{score} between the stored momen-
tary image and a reference momentary image; and

-24-

25 c) wherein the identification means comprises means for determining the identity or non-identity of the individual by evaluating the ^{second} computed first and second similarity ~~states~~.

10. Apparatus according to claim 6, wherein the defining means comprises:

5 a) a microphone for developing signals corresponding to the utterance of the predetermined key word by the individual;

Q b) a preamplifier connected to amplify the signals developed by the microphone, and

10 c) an extractor which is coupled to the pre-amplifier to define the sequence of features from the amplified signals.

11. Apparatus according to claim 6, wherein the momentary image forming means comprises:

5 a) an electronic camera for developing scanning signals corresponding to a sequence of momentary images of the voice-utterance varying portion of the individual's face;

10 b) a detector being connected to the electronic camera for converting the signals developed by the electronic camera into digital signals;

15 c) a memory coupled to the detector for receiving the digital signals; and

20 d) a key event detecting unit being connected to control the memory so that the memory stores the digital signals of the momentary image in the sequence of momentary images that corresponds to the moment of occurrence of the key event.

12. Apparatus according to claim 6, wherein the identification means comprises:

a) a first buffer for storing a reference sequence of features;

5

b) a first pattern matcher coupled to the defining means and to the first buffer for computing a first similarity ^{score} rate between the sequence of features defined by the defining means and the reference sequence of features;

10

c) a second buffer for storing a reference momentary image;

15

d) a second pattern matcher coupled to the momentary image forming means and the second buffer for computing a second similarity ^{score} rate between the momentary image formed by the momentary image forming means and the reference momentary image;

20

e) a third buffer, coupled to communicate with the momentary image forming means, for storing a reference set of parameters used by the

25 momentary image forming means to define
the occurrence of a key event;

30 f) a microprocessor control unit for
storing the reference sequence of features,
the reference momentary image and the
reference set of parameters and coupled to
communicate this stored information respec-
tively to the first, second and third buffers;
and also coupled to the pattern matchers for
determining the identity or non-identity of
the individual by evaluating the computed
first and second similarity ^{scores} rates; and

35 g) communicating means associated with the
40 microprocessor control unit for communicating
the predetermined key word to the individual
in response to a request.

13. Apparatus according to claim 6, wherein the
momentary image forming means includes a key event
detecting unit to detect the occurrence of the key
event in the utterance of the key word, the key
5 event detecting unit comprising:
a) an integrator coupled to receive the
sequence of features from the defining means
in the form of an electro-acoustical signal
defined in response to the utterance of a key
word by the individual and serving to form
10 a time-varying signal which is a function of
the amplitude of the electro-acoustical signal;

-27-

b) a control unit connected to the integrator
for detecting the beginning of the key word by
15 analyzing the time-varying signal formed by
the integrator;

c) a slope detector coupled to the control unit
and the integrator to receive the time-dependent
20 signal for detecting characteristic slope
features; and

d) a comparator coupled to the slope detector
and to the control unit for comparing the char-
acteristic slope features detected by the slope
25 detector with predetermined reference slope
features defining the key event; and

e) means coupled to the comparator to store
30 the momentary image of a voice-utterance varying
portion of the individual's face corresponding
to the key event in the utterance of the
predetermined key word when coincidence
35 between the detected slope features and the
reference slope features is determined by
the comparator.

Fetherstonhaugh & Co.,
Ottawa, Canada
Patent Agents

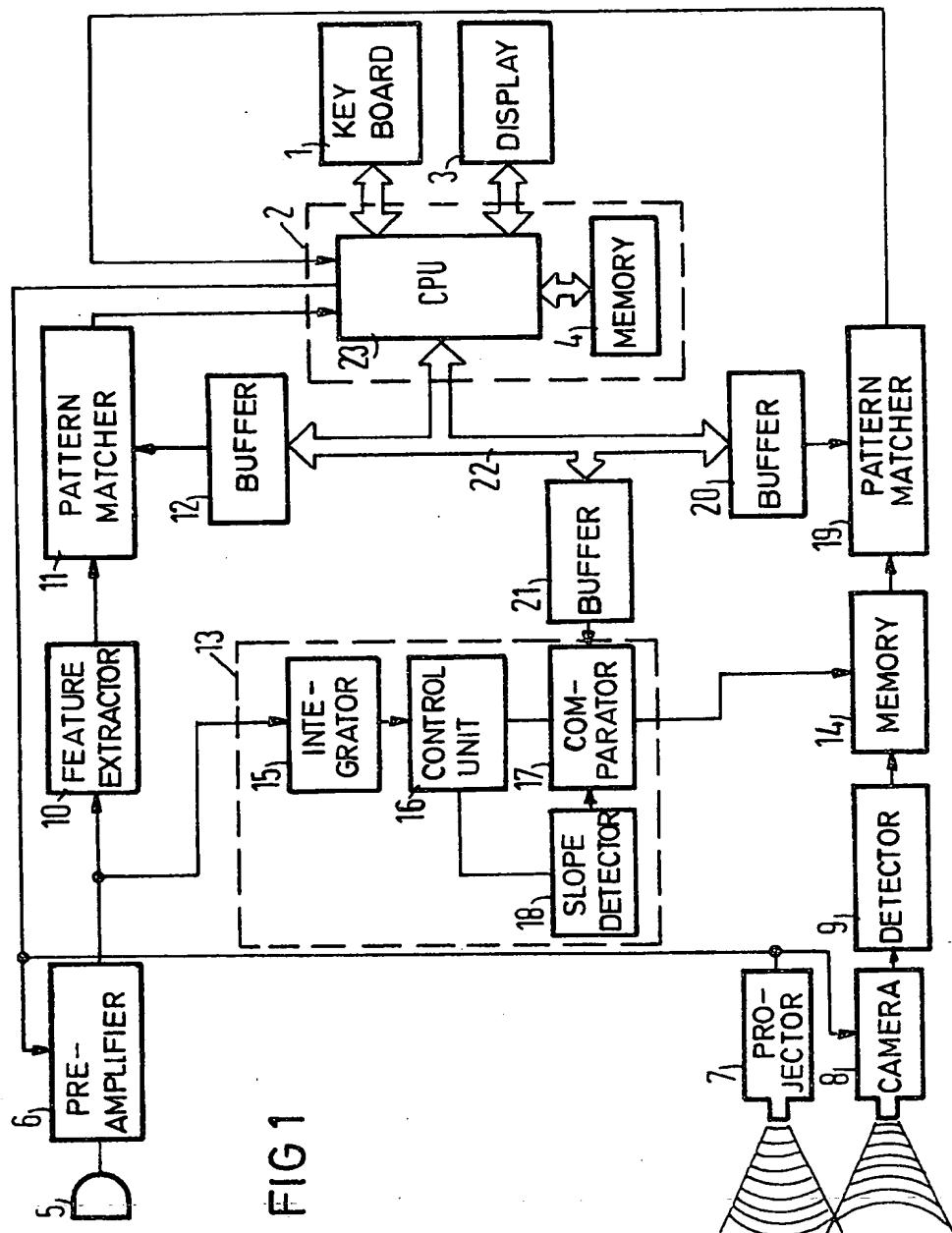


FIG 1

Patent Agents
Fechter & Hough Co.

1181856

2-2

FIG 2

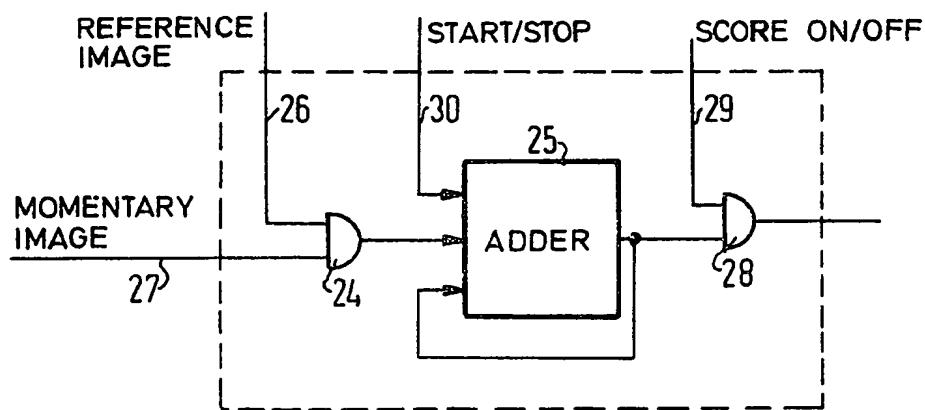
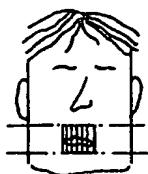


FIG 3



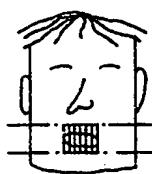
TIME EVENT

FIG 4



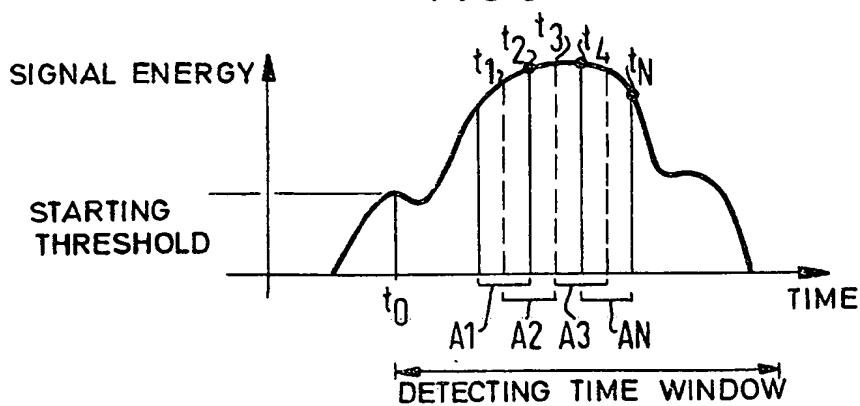
TIME EVENT

FIG 5



TIME EVENT

FIG 6



Patent Agents
Fetherstonhaugh & Co